Ross Ihaka

Ross Ihaka was born in South Auckland and received his education in a variety of small country schools in the Cook Islands and the northern North Island. He completed his BSc(Hons) and MSc Mathematics degrees at the University of Auckland before undertaking his doctoral study in statistical aspects of seismology at the University of California at Berkeley. After completing his PhD he held positions at Yale University and the Massachusetts Institute of Technology before taking up a joint position as lecturer for the departments of Computer Science, Mathematics and Statistics at the University of Auckland. Ross retired from the University of Auckland as an Associate Professor in 2017.

Ross's first experience with computing was in an Applied Mathematics course taught by Garry Tee, John Butcher and F D K Roberts. His next serious exposure to computing was in a group of Berkeley students working on a local variant of the Princeton ISP system under the guidance of Jim Reeds. He also learned about computational statistics and graphics under the guidance of Dan Asimov, Leo Briemann, David Brillinger, Louis Jaekal and Jim Reeds.

During his student career, Ross honed his computing skills by developing two of his own interactive systems for statistical computing and added a further two soon after graduating (with some input from John Hartigan). At the University of Auckland he developed the R statistical system with Robert Gentleman. Initially R was a platform for computing experiments, but it was given a superficial veneer to make it look like the Bell Laboratories' S system. The rationale for imposing this similarity was to get access to the large amount of 'free' code available for S. The R system has gone on to achieve significant success and is now widely used all over the world.

www.stats.auckland.ac.nz/ihaka-lectures

2019 IHAKA LECTURES

13 March – 3 April 2019

Rise of the machine learners: Statistical learning in the computational era

THE UNIVERSITY OF AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

SCIENCE
DEPARTMENT OF STATISTICS

2019 IHAKA LECTURES

The Ihaka Lecture Series is named after Ross Ihaka who retired from the University of Auckland as an Associate Professor in 2017. Ross, along with Robert Gentleman, co-created R – a statistical programming language now used by the majority of the world's practicing statisticians. It is hard to over-emphasise the importance of Ross's contribution to our field. We named this lecture series in his honour to recognise his work and contributions to our field in perpetuity.

Our 2019 series has the theme *'Rise of the machine learners: statistical learning in the computational era'* and will run from 13 March - 3 April.

Rise of the machine learners: Statistical learning in the computational era

Whether labelled as machine learning, predictive algorithms, statistical learning, or AI, the ability of computers to make real-world decisions is rising every year.

The 2019 Ihaka Lecture Series brings together four experts at the interface of statistics and computer science to discuss how computers do it, and how much we should let them.

Lectures commence at 6.30pm, Wednesdays Lecture Theatre PLT1, Ground Floor, Building 303, 38 Princes Street

Refreshments will be available before each lecture at 6pm in the basement foyer, Building 303.

Lecture 1: 13 March

Open source Machine Learning @ Waikato

Professor Bernhard Pfahringer
University of Waikato

The Machine Learning research group at Waikato has developed a number of open-source Machine Learning software suites, with Weka being the most prominent.

In this talk Professor Bernhard Pfahringer will introduce some of these systems, reflect on their design decisions and issues, and try to position them in the current international Machine Learning landscape.

Bernhard Pfahringer received his PhD degree from the University of Technology in Vienna, Austria, in 1995. He is a professor with the Department of Computer Science at the University of Waikato. His interests span a range of data mining and machine learning sub-fields, with a focus on streaming, randomisation and complex data.

Lecture 2: 20 March

Machine Learning with TensorFlow and R

J.J. Allaire
RStudio

TensorFlow is an open-source software library for numerical computation and machine intelligence developed by researchers and engineers working on the Google Brain Team.

In this talk J.J. Allaire will cover the R interface to TensorFlow, a suite of packages that provide high-level interfaces to both deep learning models (Keras) and standard regression and classification models (Estimators), as well as tools for cloud training, experiment management and production deployment.

The talk will also discuss deep learning more broadly (what it is, how it works, and where it might be relevant to users of R in the years ahead).

J.J. Allaire is the founder and CEO of RStudio, and the creator of the RStudio IDE. J.J. is the author of the R interfaces to TensorFlow and Keras.

Lecture 3: 27 March

Algorithmic fairness: Examples from predictive models for criminal justice

Dr Kristian Lum
Human Rights Data Analysis Group (HRDAG)

Predictive models are increasingly used in the criminal justice system to predict who will commit crime in the future and where crimes will occur. Because decisions influenced by models in this setting impact individuals' liberty, it is of the utmost importance that predictions generated by the models be 'fair'.

Using examples from predictive policing and recidivism risk assessment, Dr Kristian Lum will demonstrate how – if considerations of fairness and bias are not explicitly accounted for – such models could perpetuate and, under some circumstances, amplify undesirable historical biases encoded in the data.

Dr Lum will then give a brief overview of several notions of fairness that have been proposed in the 'algorithmic fairness' literature as solutions to these problems. She will close with a discussion of the ways in which policy, rather than data science, influence the development of these models and some alternative non-algorithmic solutions to the underlying problems these models seek to address.

Dr Kristian Lum is Lead Statistician at the Human Rights Data Analysis Group (HRDAG), where she leads the HRDAG project on criminal justice in the United States.

Dr Lum's research primarily focuses on examining the uses of machine learning in the criminal justice system and has concretely demonstrated the potential for machine learning-based predictive policing models to reinforce and, in some cases, amplify historical racial biases in law enforcement.

Lecture 4: 3 April

Statistical learning and sparsity

Professor Robert Tibshirani
Stanford University

In this talk Professor Robert Tibshirani will review the lasso method for high dimensional supervised learning and discuss some new developments in the area, including the Pliable Lasso, and post-selection inference for understanding the important features.

Professor Tibshirani will also describe some applications of these methods to his own collaborative work, including prediction of platelet usage at Stanford Hospital.

Rob Tibshirani is a professor of statistics and biomedical data science at Stanford University. His main interests are in applied statistics, biostatistics and data science. He is most well-known for the LASSO, which is a shrinkage and selection method for linear regression. He is the co-author of the books *Generalized Additive Models* (with T. Hastie), *An Introduction to the Bootstrap* (with B. Efron), *An Introduction to Statistical Learning* (with G. James, D. Witten and T. Hastie), *Sparsity in Statistics* (with T. Hastie and M. Wainwright), and the widely used *Elements of Statistical Learning* (with T. Hastie and J. Friedman). His current research focuses on problems in biology and genomics, medicine, and industry.